Bi-Event Subtraction Technique at Hadron Colliders

Bhaskar Dutta[1], Teruki Kamon[1,2], Nikolay Kolev[3], and Abram Krislock[1]

[1] *Department of Physics & Astronomy, Mitchell Institute for Fundamental Physics,*
Texas A&M University, College Station, TX 77843, USA

[2] *Department of Physics, Kyungpook National University, Daegu 702-701, South Korea and*

[3] *Department of Physics, University of Regina, Regina, SK S4S 0A2, Canada*

We propose the Bi-Event Subtraction Technique (BEST) as a method of modeling and subtracting large portions of the combinatoric background during reconstruction of particle decay chains at hadron colliders. The combinatoric background arises when it is impossible to know experimentally which observed particles come from the decay chain of interest. The background shape can be modeled by combining observed particles from different collision events and be subtracted away, greatly reducing the overall background. This idea has been demonstrated in various experiments in the past. We generalize it by showing how to apply BEST multiple times in a row to fully reconstruct a cascade decay. We show the power of BEST with two simulated examples of its application towards reconstruction of the top quark and a supersymmetric decay chain at the Large Hadron Collider.

The Large Hadron Collider (LHC) is up and running since 2009. Many models of particle physics beyond the Standard Model (SM) predict new particles which can be tested at the LHC. Heavy colored objects are expected to be produced at the LHC, followed by a chain of subsequent decays, according to such new models. Thus, we must fully or partially reconstruct these cascade decays from the particles which can be detected. However, reconstructions of these decays become experimentally difficult because it is impossible to know which particles come from the cascade decay we wish to reconstruct. The inevitable inclusion of particles which do not come from the cascade decay of interest is referred to as combinatoric background.

This combinatoric background can be removed easily in some cases by powerful subtraction techniques. For instance, the Z boson can decay into oppositely charged, same flavored leptons: $Z \rightarrow e^+e^-/\mu^+\mu^-$. Leptons are easy to detect in the collider setting, and their charges can easily be measured. To reconstruct the Z boson from these leptons, it is easy to collect a sample of Opposite-Sign Same-Flavor (OSSF) lepton pairs and construct the dilepton invariant mass for each pair. To model the combinatoric background, a sample of Opposite-Sign Opposite-Flavor (OSOF) lepton pairs is selected as well. These OSOF lepton pairs cannot possibly both come from a single Z boson, and so they model the combinatoric background well. Performing the OSSF$-$OSOF subtraction of the invariant mass distributions (possibly using some normalization factor c), $h^{\text{OSSF}-\text{OSOF}}(m_{\ell\ell}) = h^{\text{OSSF}}(m_{\ell\ell}) - ch^{\text{OSOF}}(m_{\ell\ell})$, yields a distribution which shows a clear peak of the Z boson mass.

However, such subtraction techniques are not available for jets, whose charges and flavors cannot so easily be determined. Thus, we introduce the Bi-Event Subtraction Technique (BEST) in which the combinatoric background of jets is modeled by combining jet information from a different event (or bi-event). This technique of modeling the combinatoric background by combining information from different events has been used before [1]. However, here we generalize it, by applying it to jets. Moreover, we have shown that it can be used multiple times for the same decay chain reconstruction.

The basic idea of BEST can be demonstrated for the reconstruction of the W boson decaying into two jets. For this case, a signal may be seen if a sample of jet pairs is collected for each event to construct the dijet invariant mass distribution, $h^{\text{same}}(m_{jj})$. Here, the "same" suggests that the jet pairs come from the same event. Some of the jet pairs in the same event distribution may come from a single W boson decay in the events, while other jet pairs will be combinatoric background. By collecting another sample of jet pairs where each jet comes from a *different* event, the bi-event distribution, $h^{\text{bi}}(m_{jj})$, can be formed. This bi-event distribution will have no jet pairs which come from a single W boson. Thus, this bi-event distribution models a large amount of the combinatoric background well. The $h^{\text{bi}}(m_{jj})$ distribution can be normalized to the $h^{\text{same}}(m_{jj})$ distribution in the region of pure background (well away from the W boson mass peak). For instance, the normalization factor can be calculated as

$$
C_{jj}^{\text{BEST}} = \frac{\displaystyle\int_{150\ \text{GeV}}^{500\ \text{GeV}} h^{\text{same}}(m_{jj})\mathrm{d}m_{jj}}{\displaystyle\int_{150\ \text{GeV}}^{500\ \text{GeV}} h^{\text{bi}}(m_{jj})\mathrm{d}m_{jj}}. \tag{1}
$$

This normalization factor can be used when the shapes of these distributions are very close in this region. If the shapes of these distributions are not close, it could be due to some new physics. For instance, and additional resonance in the $h^{\text{same}}(m_{jj})$ distribution could cause a mismatch in the shapes. However, it would be easy enough to recalculate the normalization taking an overall range which excludes the additional resonance. It should be noted that one needs a detailed systematic study of the shape from different physics processes. This is beyond the scope of this paper.

Finally, the BEST is performed:

$$
h^{\text{BEST}}(m_{jj}) = h^{\text{same}}(m_{jj}) - C_{jj}^{\text{BEST}} h^{\text{bi}}(m_{jj}). \tag{2}
$$

The resulting dijet distribution shows a W boson mass peak with most of the combinatoric background removed.

If we wish to reconstruct decay chains involving these W bosons, we can take BEST even further. For instance, we can completely reconstruct the top quark from the decay chain $t \rightarrow bW \rightarrow bjj$. We can apply BEST again while combining the b jets with the reconstructed W bosons in order to reconstruct the top quark. However, this requires a more general application of BEST than has been used before.

For this example, we will refer to the same-event histograms by denoting the jets in the subscript as j and

b for jets and b-jets respectively. For the bi-event histograms, we denote the jets in the subscript as j' and b'. [3]
Thus we now denote our histograms and normalization factor from Eqs. (1) and (2) as:

$$h^{\text{same}}(m_{jj}) \equiv h_{jj}(M_{jj}), \tag{3a}$$

$$h^{\text{bi}}(m_{jj}) \equiv h_{jj'}(M_{jj}), \tag{3b}$$

$$C_{jj}^{\text{BEST}} \equiv C_{jj}^{\text{BEST\#1}}, \tag{3c}$$

$$h^{\text{BEST}}(m_{jj}) \equiv h_{jj}^{\text{BEST\#1}}(m_{jj}) \tag{3d}$$

To combine the reconstructed W bosons with the b-jets to reconstruct the top quarks, we will need the following four additional histograms in order to perform two applications of BEST: $h_{bjj}(m_{bjj})$, $h_{bjj'}(m_{bjj})$, $h_{b'jj}(m_{bjj})$, and $h_{b'jj'}(m_{bjj})$. We perform the first BEST using the normalization factor calculated above in Eq. (1):

$$h_{bjj}^{\text{BEST\#1}}(m_{bjj}) = h_{bjj}(m_{bjj}) - C_{jj}^{\text{BEST\#1}} h_{bjj'}(m_{bjj}) \tag{4a}$$

$$h_{b'jj}^{\text{BEST\#1}}(m_{bjj}) = h_{b'jj}(m_{bjj}) - C_{jj}^{\text{BEST\#1}} h_{b'jj'}(m_{bjj}) \tag{4b}$$

Next we calculate another normalization factor for the second BEST which involves the combinatoric background of the b-jets. Once again, the range of this normalization factor is aimed at the region of pure background away from the top quark mass peak. Thus, it is calculated as:

$$C_{bjj}^{\text{BEST\#2}} = \frac{\int_{200\ \text{GeV}}^{500\ \text{GeV}} h_{bjj}^{\text{BEST\#1}}(m_{bjj}) \mathrm{d}m_{bjj}}{\int_{200\ \text{GeV}}^{500\ \text{GeV}} h_{b'jj}^{\text{BEST\#1}}(m_{bjj}) \mathrm{d}m_{bjj}}. \tag{5}$$

With this normalization factor, we can finally perform the second BEST:

$$h_{bjj}^{\text{BEST\#2}}(m_{bjj}) = h_{bjj}^{\text{BEST\#1}}(m_{bjj}) - C_{bjj}^{\text{BEST\#2}} h_{b'jj}^{\text{BEST\#1}}(m_{bjj}). \tag{6}$$

Here, the resulting histogram will show a clean top quark mass peak with most of the combinatoric background removed. To clean up the resulting distribution even more, other subtraction techniques can also be employed, such as a sideband subtraction for the W boson reconstruction. Each additional subtraction will double the number of initial histograms which are needed for all of the subtractions.

We demonstrate this powerful technique by using it to extract $W \to jj$ for (i) $t\bar{t}$ events at $\sqrt{s} = 7$ TeV and (ii) SUSY events at $\sqrt{s} = 14$ TeV within LHC simulations.

For the $t\bar{t}$ events, we generate hard scattering LHC collision events using `ALPGEN` [2], perform the cascade decays with `PYTHIA` [3], and perform a LHC detector simulation using `PGS4` [4]. The W+jets events are the main source of background for finding the top quark, so we generate these events in the same way. This background is mixed in randomly, according to production cross-sections, with our $t\bar{t}$ events. After `PGS4` is finished with these events, we select events for analysis with the following cuts [5]: (i) Number of leptons, $N_\ell = 1$, where $p_T^{(\ell)} \geq 20$ GeV and $p_{T,iso}^{(\ell)} \leq 0.1 \times p_T^{(\ell)}$; (ii) Missing transverse energy, $\not{E}_T \geq 20$ GeV; (iii) Number of jets, $N_j \geq 3$, where $p_T^{(j)} \geq 30$ GeV and at least one jet has been tightly b-tagged [4]; (iv) Number of taus, $N_\tau = 0$ for taus with $p_T^{(\tau)} \geq 20$ GeV [4].

With our events selected in this way, we pair up jets (which are not b-tagged) to fill the same-event and bi-event $h(m_{jj})$ distributions as described above. Each jet pair must have $\Delta R \geq 0.4$. To fill the bi-event distribution, we refer to jets from the previous event which has passed the same cuts as listed above. Once the distributions are filled with all events, we normalize the shape of the $h^{\text{bi}}(m_{jj})$ distribution as described by Eq. (1). Then we perform our BEST. The result of this subtraction can be seen in Fig. 1, which shows a drastic reduction in the background obscuring the W boson reconstruction. Note that the bi-event distribution models the combinatoric background of any jet pairs which are not correlated by decay chains or event kinematics. Thus, BEST in this case removes (i) the combinatoric background from events with W bosons (coming from t decays) and (ii) uncorrelated jet pairs coming from our W+jets background sample.
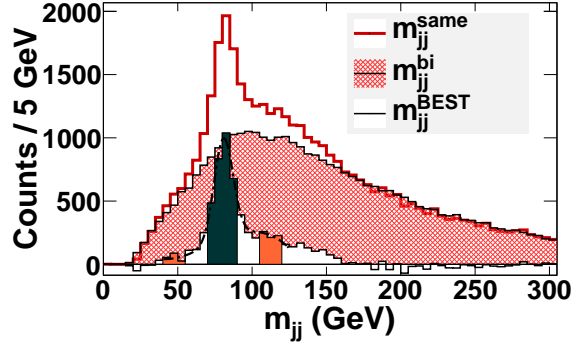
FIG. 1: The dijet invariant mass distribution, m_{jj}. This plot shows the same-event (m_{jj}^{same}), bi-event (m_{jj}^{bi}), and BEST (m_{jj}^{BEST}) distributions as described in the text. The BEST distribution is fitted with a gaussian plus cubic function, to find the W boson mass peak and surrounding background. The BEST distribution is also split up into regions for a sideband subtraction used for reconstructing an invariant mass between a W boson and a b tagged jet. The W region is dark cyan filled, while the sidebands are orange filled. For an integrated luminosity of 2 fb^{-1}, we find the W boson mass, $m_W = 81.11 \pm 0.32$ GeV.
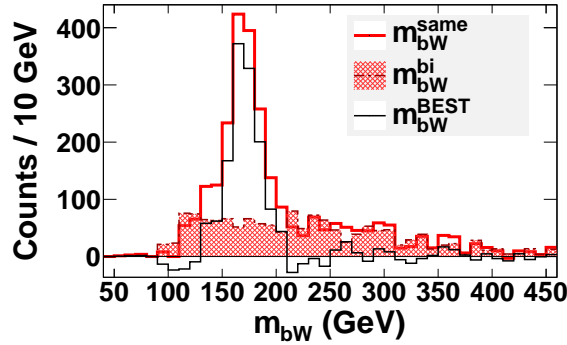

FIG. 2: The W plus b invariant mass distribution, m_{bW}. This plot shows the same-event, bi-event, and BEST distributions as described in the text. For an integrated luminosity of 2 fb^{-1}, we find the top quark mass, $m_t = 170.5 \pm 1.5$ GeV. The top quark mass is set within `ALPGEN` as $m_t = 174.3$ GeV.

Once we have found the W boson with this first application of BEST, we can combine the W boson with a b-jet to find the top quark. To remove additional background from the W signal, we perform a sideband subtraction. To do this, we split up the dijet signal into a W boson mass region, where 70 GeV $\leq m_{jj} \leq$ 90 GeV, and two sideband regions, 40 GeV $\leq m_{jj} \leq$ 55 GeV and 105 GeV $\leq m_{jj} \leq$ 120 GeV. We form the dijet (W) plus b invariant mass, keeping track of whether the dijet system was in the W window or sideband windows. In this way we make the W band ($h_{W\text{band, BEST}}(m_{bW})$) and sideband ($h^{\text{SB, BEST}}(m_{bW})$) distributions. The sideband distribution models the remaining background of W's very well. By fitting the $h^{\text{BEST}}(m_{jj})$ by a gaussian function, $f(m_{jj}^{\text{BEST}})$, plus a background function, $g^{\text{BG}}(m_{jj}^{\text{BEST}})$, we can find the shape of the background distribution which remains. Then we calculate a normalization factor:

$$C_{jj}^{\text{SB}} = \frac{\int_{W\text{band}} g^{\text{BG}}(m_{jj}^{\text{BEST}}) \mathrm{d}m_{jj}^{\text{BEST}}}{\int^{\text{SBs}} g^{\text{BG}}(m_{jj}^{\text{BEST}}) \mathrm{d}m_{jj}^{\text{BEST}}}, \tag{7}$$

Using this normalization factor, we perform the sideband subtraction,

$$\begin{aligned} h^{\text{SBsub, BEST}}(m_{bW}) = \\ h_{W\text{band, BEST}}(m_{bW}) - C_{jj}^{\text{SB}} h^{\text{SB, BEST}}(m_{bW}). \end{aligned} \tag{8}$$
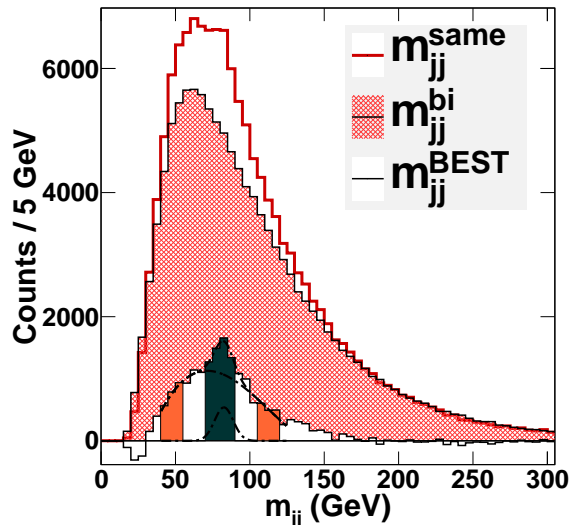
FIG. 3: The dijet invariant mass distribution, m_{jj} from our nuSUGRA events mixed with SM backgrounds. The BEST has already been performed. The BEST distribution is fitted and split up into regions for a sideband subtraction used for reconstructing an invariant mass between a W boson and a leading jet. The W region is dark cyan filled, while the sidebands are orange filled. Here we find the W boson mass, $m_W = 82.4 \pm 1.0$ GeV. This plot is for an integrated luminosity of 100 fb^{-1}.

This subtraction removes even more of the W combinatoric background.

Lastly, to remove the combinatoric background of b-jets, we can perform our BEST again. We form the $h^{\text{SBsub, BEST}}(m_{bW})$ distribution again, this time using b jets which come from a different event as the W. Again, this models the combinatoric background very well, since the W and b from different events cannot possibly come from a single top quark. We can calculate a normalization factor as before analogous to Eq. (1) in the range 200 GeV $\leq m_{bW} \leq$ 500 GeV (a little away from the top mass peak). Using this normalization factor, we can perform the final BEST, analogously to that shown in Eq. (6). The resulting m_{bW} distribution after this last application of BEST is shown in Fig. 2, which shows a very clean looking top peak.

In the context of top reconstruction, other groups have come up with some techniques to eliminate the combinatoric background. In experimental top reconstruction [5–7], combinatoric background is eliminated by assuming a very particular event topology. By selecting certian events, the combinatoric background is eliminated by essentially choosing the jet combinations which form the best W and t masses. For SM $t\bar{t}$ events, this reconstruction works quite well to measure the top mass. However, these methods cannot be employed to reconstruct t quarks from beyond SM sources. On the other hand, some phenomenological studies of top production from beyond SM use top tagging [8] to identify the top correctly, and thus reduce the combinatoric background. However, the top tagging relies on the production of a boosted top from the decay of a heavy new particle. Also, although this top tagger has a large efficiency, it seems the fake rate from SM backgrounds may be large. These experimental and phenomenological techniques may be more precise than BEST (although, a thorough study would be needed to compare them). However, the advantage of BEST is that it does not require any assumptions about the event topology or having boosted tops.

This example from the SM shows the power of BEST. Additionally, BEST is useful for searches and measurements of models beyond the SM. Thus, we also demonstrate the use of BEST for a supersymmetry (SUSY) model. The model we choose is the non-universal generalization of the minimal supergravity model [9] i.e., nuSUGRA. In this nuSUGRA model, the Higgs masses are not unified with the other scalar masses at the grand unified scale. This allows for a more general mass spectrum than that of the mSUGRA model. The indication of the preference for the nuSUGRA model at the LHC is that the neutralino masses will not have the mass ratios predicted by mSUGRA. This nuSUGRA model can also predict the correct amount of dark matter in the universe today. In particular, a large parameter space region of this model has an abundance of W bosons being produced [10]. These W bosons must be found and utilized to reconstruct the model. Thus, this is a useful model to examine with BEST.
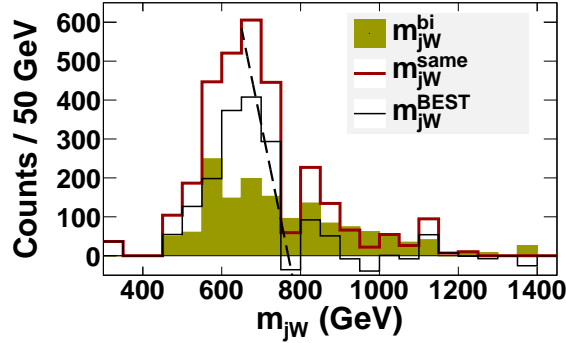
FIG. 4: The W plus jet invariant mass distribution, m_{jW}. This plot shows the same-event, bi-event, and BEST distributions as described in the text. BEST removes the background obscuring the endpoint. For an integrated luminosity of 100 fb^{-1}, we find the endpoint to be 769 ± 18 GeV. This is within 2σ of the theoretical endpoint, which is 738.8 GeV for the most probable decay chain of this type, $\tilde{q} \rightarrow q + \tilde{\chi}_4^0 \rightarrow q + W^\pm + \tilde{\chi}_1^\mp$.

We choose a benchmark point for the nuSUGRA model for this demonstration: $m_0 = 360$ GeV, $m_{1/2} = 500$ GeV, $\tan\beta = 40$, $A_0 = 0$, and $m_H = 732$ GeV, with the top mass set as $m_t = 172.6$ GeV. This point in parameter space predicts an abundance of W bosons at the LHC due to neutralino or chargino decays. The decay chain we wish to partially reconstruct is: $\tilde{q} \rightarrow q + \tilde{\chi}_1^\pm \left(\tilde{\chi}_4^0 \right) \rightarrow q + W^\pm + \tilde{\chi}_1^0 \left(\tilde{\chi}_1^\mp \right)$. Our BEST has been used to analyze this signal already, with the details shown in [10].

To simulate events for this demonstration, we once again use PYTHIA and PGS4. The SUSY mass spectrum is generated using ISAJET [11]. We also use ALPGEN to simulate some SM backgrounds. The primary SM backgrounds for the events we wish to analyze are Z+jets, W+jets, and $t\bar{t}$ events. We mix these SM backgrounds in randomly with our SUSY signal events.

To help reduce the SM backgrounds, we use the following selection cuts, which are refined from the cuts in [10]: (i) Missing transverse energy, $\not{E}_T \geq 180$ GeV; (ii) Number of jets, $N_j \geq 4$, where $p_T^{(j)} \geq 30$ GeV; (iii) Minimum $\Delta\phi$ between leading three jets and missing transverse energy, $\Delta\phi^{\min} \geq 0.5$; (iv) Leading jet transverse momenta, $p_T^{(1\text{st } j)} \geq 300$ GeV and $p_T^{(2\text{nd } j)} \geq 200$ GeV; (v) ΔR between leading jets, $\Delta R(1\text{st } j, 2\text{nd } j) \leq 3.2$; (vi) Scalar sum, $p_T^{(1\text{st } j)} + p_T^{(2\text{nd } j)} + 3 \cdot \not{E}_T \geq 1600$ GeV.

With these event selection cuts, we begin to pair up the sub-leading jets as we did for the $t\bar{t}$ analysis, perform the BEST to find the W bosons, then combine the W's with the leading jets to reconstruct the desired decay chain. While pairing up the jets, we use the additional cut $0.4 \leq \Delta R(jj) \leq 1.5$. We once again perform a sideband subtraction to help clean up any excess background involved with finding the W bosons. When combining the W candidates (jet pairs) with leading jets, we keep only those combinations where $\Delta R(W, j) \geq 1.0$. We use BEST again on the leading jet as well, to remove combinatoric background from the leading jets which are not from our desired decay chain. The result of this analysis can be seen in Figs. 3 and 4. Note in Fig. 3 that the W boson mass peak can barely be seen in the same-event histogram, but is clearly visible after the application of BEST.

In conclusion, BEST is a powerful subtraction technique which can find and reconstruct particles normally hidden by the combinatoric background, as shown in Fig. 3. It is useful for the further understanding of the SM as well as models beyond the SM. It can be utilized without information about the charge or flavor of the particles involved. BEST can, therefore, improve any current and future collider study and help us detect new particles, measure their masses and determine model parameters accurately.

Acknowledgements

and S. Su for useful discussions, and Y. Oh for providing computer support.

[1] M.G. Albrow *et al.*, Nucl. Phys. B **114** (1976) 365; The Belle Collaboration, Phys. Rev. D **77** (2008) 112003, and Phys. Rev. D **78** (2008) 032011; ATLAS Collaboration, N. Ozturk, [arXiv:hep-ph/0710.4546]; The CMS Collaboration, J. High Energy Phys **09** (2010) 091.

[2] M. Mangago, M. Moretti, F. Piccinini, R. Pittau, and A. Polosa, J. High Energy Phys **07** (2003) 001. We use `ALPGEN` version 2.13.

[3] T. Sjostrand, S. Mrenna, and P. Skands, J. High Energy Phys. **05** (2006) 026. We use `PYTHIA` version 6.411 with TAUOLA.

[4] `PGS4` is a parameterized detector simulator. We use version 4 (`http://www.physics.ucdavis.edu/~conway/research/software/pgs/pgs4-general.htm`) in the CMS detector configuration. We assume the τ identification efficiency with $p_{\mathrm{T}}^{\mathrm{vis}} > 20$ GeV is 50%, while the probability for a jet being mis-identified as a τ is 1%. The b-jet tagging efficiency in PGS is $\sim$42% for $E_{\mathrm{T}} > 50$ GeV and $|\eta| < 1.0$, and degrading between $1.0 < |\eta| < 1.5$. The b-tagging fake rate for c and light quarks/gluons is $\sim$ 9% and 2%, respectively.

[5] CDF Collaboration, Phys. Rev. Lett. **105** (2010) 252001.

[6] The CMS Collaboration, CMS-TOP-10-002, CMS-PAS-TOP-10-009.

[7] The ATLAS Collaboration, ATLAS-CONF-2011-033.

[8] T. Plehn, M. Spannowsky, M. Takeuchi, and D. Zerwas, JHEP 1010 (2010) 078; T. Plehn, G. P. Salam, and M. Spannowsky, Phys. Rev. Lett. 104, 111801 (2010), 0910.5472; `http://www.thphys.uni-heidelberg.de/~plehn/heptoptagger/`

[9] D. Z. Freedman, P. Van Niewenhuisen, and S. Ferrara, Phys. Rev. D **13** (1976) 3214; S. Deser and B. Zumino, Phys. Lett. B **65** (1976) 369; A.H. Chamseddine, R. Arnowitt, and P. Nath, Phys. Rev. Lett. **49** (1982) 970; R. Barbieri, S. Ferrara, and C.A. Savoy, Phys. Lett. B **119** (1982) 343; L. Hall, J. Lykken, and S. Weinberg, Phys. Rev. D **27** (1983) 2359; P. Nath, R. Arnowitt, and A.H. Chamseddine, Nucl. Phys. B **227** (1983) 121; For a review, see P. Nilles, Phys. Rep. **100** (1984) 1.

[10] B. Dutta, T. Kamon, A. Krislock, N. Kolev, and Y. Oh, Phys. Rev. D **82** (2010) 115009.

[11] F. E. Paige *et al.*, [arXiv:hep-ph/0312045]. We use `ISAJET`version 7.74.